Technical Report 2010-002
Critical region analysis of scalar fields in arbitrary dimensions

Madjid Allili[a], Marc Ethier[b], and Tomasz Kaczynski[b]

[a]Bishop's University, 2600 College St., Sherbrooke, Quebec, J1M 1Z7, Canada;
[b]Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke, Quebec, J1K 2R1, Canada

ABSTRACT

The exploration of multidimensional scalar fields is commonly based on the knowledge of the topology of their isosurfaces. The latter is established through the analysis of critical regions of the studied fields. A new method, based on *homology theory*, for the detection and classification of critical regions in multidimensional scalar fields is proposed in this paper. The use of computational homology provides an efficient and successful algorithm that works in all dimensions and allows to generalize visual classification techniques based solely on the notion of connectedness which appears insufficient in higher dimensions. We present the algorithm, discuss details of its implementation, and illustrate it by experimentations in two, three, and four dimensional spaces.

Keywords: Height fields, critical regions, isosurfaces, homology, Morse index, Conley index

1. INTRODUCTION

The exploration and visualization of scalar fields are commonly based on the study of their isosurfaces for given isovalues. By varying the isovalues and studying each corresponding isosurface, one can completely explore a given scalar field, at any desired level of detail. Moreover, in domains such as medical image processing, many imaging modalities give images (such as CT and MRI scan data images) that are piecewise constant, because the intensity is related to tissue type. It follows that isosurfaces correspond approximately to tissue regions.

This approach requires solving two difficult problems. The first problem consists of extracting a topologically correct isosurface for each given isovalue. The second issue is about the impossibility of undertaking the study of the different representations of all possible isosurfaces in large scale data sets. Therefore, one needs to limit the study to isovalues where "interesting" behavior occurs. The analysis of topological properties of a given discrete scalar field makes it possible to determine critical regions with interesting isosurface behavior. By determining the locations and nature of critical components of the data, one can track all the fundamental changes and reconstruct a topologically correct isosurface for any given isovalue using meshing techniques and marching cubes algorithms (see[1] and the references therein).

The classical approach to detect and classify critical points is the Morse theory[2,3] that relies on strong smoothness and non-degeneracy assumptions, which appear not realistic for discrete models. Among the non-degeneracy assumptions, it is required that critical points be isolated and have distinct critical values. This assumption fails when considering, for example, a height field of a terrain with water surfaces as minimum plateaus, sandbars near a seashore which may become saddle isolines at a low tide, and ridges of volcano craters as maximum closed isolines.

Classical Morse theory can be applied directly in the analysis of discrete data by virtually forcing the data to obey the assumptions of the theory. This practice is standard in image processing where the discrete data is smoothed by convolution with a suitable kernel such as the Gaussian. It is mathematically validated by the argument that a small perturbation of a given function can bring it to the *generic case* where all critical points are non-degenerate, isolated, and correspond to distinct critical values. It may also be viewed as *simulating the smoothness* on the discrete data via interpolation. This approach imposes that the computed features have the same structural form as those found in the smooth data. Among most systematic studies of that kind one could point to the work of Edelsbrunner *et al.*[4,5] While this approach may be natural in some problems, there are many other situations where it is not applicable. One, addressed by Arnold,[6] arises in

Further author information:
M. Allili: E-mail: mallili@ubishops.ca
{M. Ethier; T. Kaczynski}: E-mail: {marc.ethier; tomasz.kaczynski}@usherbrooke.ca

the study of a parameterized family of equilibrium states. A small perturbation can remove a degeneracy at one parameter value but it would create a new one at another value. Another situation arises in singularity theory.[6,7] Critical regions of polynomial functions of several variables are extensively studied there, under the name of singular curves and surfaces. The *Whitney-Cayley umbrella*[?] given by

$$x^2 - y^2 z = 0$$

is a simple classical example of an algebraic level surface in R^3 self-intersecting at a singular curve (the z-axis).

Standard methods for isosurface extraction such as marching cubes perform quite well for simple non-singular level surfaces $f = c$. However, the mesh generation marching algorithms may no longer be used near the singular zones $\nabla f = 0$ which may consist of isolated critical points but they may also be formed of curves at which isosurfaces self-intersect. In particular, closed singular curves exhibiting non-trivial topology occur in many applications and are studied in this paper. In singular zones, one has to investigate the local topology nearly on a pixel level, in order to determine how to locally complete the construction of the studied surface so as to preserve its global properties.

For discrete data in two dimensions, several successful approaches to define and classify critical regions have appeared.[?,1] In that case, the detection and classification of a critical region $\mathscr{C}$ can be done by studying the number of connected components of its *superlevel set* and *sublevel set*, which are the subsets of their immediate neighborhoods where the values of the function are respectively greater and smaller than $f(\mathscr{C})$. The method given by Weber *et al.*[1] works also for classification of critical points of continuous but not necessarily smooth functions in $\mathbb{R}^3$.

However, critical components in $\mathbb{R}^3$ and critical points and components in higher dimensions require more attention and the number of connected components may no longer be sufficient to do an accurate and complete classification. We present in Sec. 2 two examples illustrating this difficulty. To overcome the problem, one has to take into account such topological invariants as homotopy type, homotopy groups, or homology groups. We focus in the present work on using homology descriptors for the classification of critical components, because there is a vast library of convenient programs computing homology such as ChomP,[?] while the homotopy type is not constructive, and the computability of homotopy groups is problematic.

In Sec. 2, we recall the basic topology notions concerning Morse and Conley indices and we give examples motivating the use of homology descriptors for the classification of critical components in higher dimensions.

The main results are presented in Sec. 3. We propose a method and the corresponding algorithms for detecting and classifying critical regions of discrete functions inspired by the Conley index theory. We give algorithmic definitions of regular and critical regions for functions defined on a unit-size cubical grid in $\mathbb{R}^d$. This choice is well justified in practice, because many applications in 3D imaging science produce regularly gridded data by sampling scalar fields at uniform intervals of time and space. This data is given in the form of scalar functions defined on the vertices of a hypercubic decomposition of space. Regularly gridded data does not require an explicit storage of cell adjacency information, which would be necessary, if we wanted to subdivide each cube into tetrahedral grids. Thus it results in lower complexity and lower storage requirements. Note that irregularly gridded data is typically organized into tetrahedral cells using techniques such as the ones based on Voronoi diagrams and Delaunay triangulations.[3] Our method is valid in all dimensions. It requires verifying whether or not relative homology groups of certain sets are trivial. In our case, the verification is implemented at a low cost, since it reduces to so-called *elementary collapses*. For this, see Sec. 2.4 of Ref. **?**.

Sec. 3 is concluded with experimentation using a computer program based on the presented algorithms. We test the code on 2D and 3D scalar fields obtained by discretizing functions given by mathematical formulas, in order to compare the program output with anticipated results.

In Sec. 4, we discuss some open questions and prospects for future applications to construction of isosurfaces.

2. TOPOLOGY BACKGROUND AND MOTIVATING EXAMPLES

Let X be a bounded domain in $\mathbb{R}^d$ and $f : X \to \mathbb{R}$ a function of class C^2. A point $p \in X$ is *critical* if the gradient ∇f vanishes at p and it is called *regular* otherwise. The function f is called *generic* or *Morse* if all of its critical points p are *non-degenerate*, that is, the Hessian $H_f(p) = \det D^2 f(p)$ is nonzero. This condition implies in particular that p is an isolated critical point.

Given a generic function f, the index of its critical point p, denoted by $\lambda(p)$, is the number of negative eigenvalues of $D^2 f(p)$. If $\lambda(p) = 0$, p is a local minimum and if $\lambda(p) = d$, it is a local maximum. The values $0 < \lambda(p) < d$ characterize saddles which may topologically differ from each other when $d > 3$.

Consider a function whose critical points are still isolated but possibly degenerate. Then there may occur critical points p of "inflection type" which are *inessential* in the sense that a perturbation of f in a small isolating neighborhood N of p may cause the removal of singularity from N. We are interested in critical points which are *essential* in the sense that the topology of the *level set* given by $f(x) = c$ changes when the parameter value c crosses $f(p)$. Alternatively, one may characterize the essentiality by means of *sublevel sets* $f(x) < c$ or *superlevel sets* $f(x) > c$. A weakness of this approach is that it is global, that is, the computation of topological invariants has to be done on the whole domain X, far away from p. We want to localise this computation to a suitably defined *isolating neighborhood* N of p. Thus we consider the localized sublevel and superlevel sets given by

$$N_n = \{x \in N \mid f(x) < f(p)\},$$
$$N_p = \{x \in N \mid f(x) > f(p)\},$$

The intersection of their closures is the level set

$$N_z = \{x \in N \setminus \{p\} \mid f(x) = f(p)\}.$$

Whether p is a regular point or an inessential critical point, if N is a disc, the sets N_n and N_p should be its topological half-discs. At a maximum or minimum point, one set is $N \setminus \{p\}$ and another one is empty. At a critical saddle such as $p = 0$ of the function $f(x,y) = x^2 - y^2$, each of the sets N_n and N_p consist of two disjoint wedges limited by the crossing lines $x^2 = y^2$. We may be tempted to use the numbers n_n, n_p and n_z of connected components of N_n, N_p, and N_z respectively, to distinguish various types of topological critical points. This approach seems to work well in dimension 2 but it fails to characterize critical points in higher dimensions, as the following example illustrates.

EXAMPLE 2.1. Let $f : \mathbb{R}^4 \to \mathbb{R}$ be given by

$$f(x,y,z,t) = x^2 + y^2 - z^2 - t^2.$$

Then $p = 0$ is an isolated non-degenerate critical point with the value 0. It is a saddle point of Morse index $\lambda(p) = 2$. Consider its neighborhood $N = \overline{B}(0,r)$ the closed ball of radius r centered at 0. Since f is radially homogeneous, the choice of r is not important so we may assume $r = 1$. For the same reason, the radial projection of $N \setminus \{0\}$ onto the unit sphere S^3 given by

$$x^2 + y^2 + z^2 + t^2 = 1$$

is a homotopy deformation. Thus, for visualization purposes, we may restrict our analysis of the discussed sets N_n, N_p and N_z to their traces on the 3–dimensional sphere S^3:

$$L_n = N_n \cap S^3 = \{(x,y,z,t) \in S^3 \mid x^2 + y^2 < z^2 + t^2\},$$
$$L_p = N_p \cap S^3 = \{(x,y,z,t) \in S^3 \mid x^2 + y^2 > z^2 + t^2\},$$

and

$$L_z = N_z \cap S^3 = \{(x,y,z,t) \in S^3 \mid x^2 + y^2 = z^2 + t^2\}.$$

Since L_z can be expressed by $x^2 + y^2 = 1/2 = z^2 + t^2$, it is a torus (product of two circles). Its complements L_n and L_p in S^3 are *toroids* that is, products of a circle by a 2-disc. All three sets are connected. The same is true for a regular point, hence the number of connected components does not permit to distinguish a saddle from a regular point in $\mathbb{R}^4$. On the other hand, one can show that each of these sets has the homotopy type of a circle, so more sophisticated topological invariants such as homology may detect the criticality.

In the study of critical regions or, more explicitly, connected regions consisting of non-isolated critical points, the three-dimensional space $\mathbb{R}^3$ is sufficient to construct examples showing that the connected components of sublevel and superlevel sets do not provide sufficient information.

EXAMPLE 2.2. Let X be a toroid in $\mathbb{R}^3$ obtained by revolution of the disc D_{r_1,r_2} given by

$$(x - r_2)^2 + z^2 \le r_1^2, \ \ y = 0$$

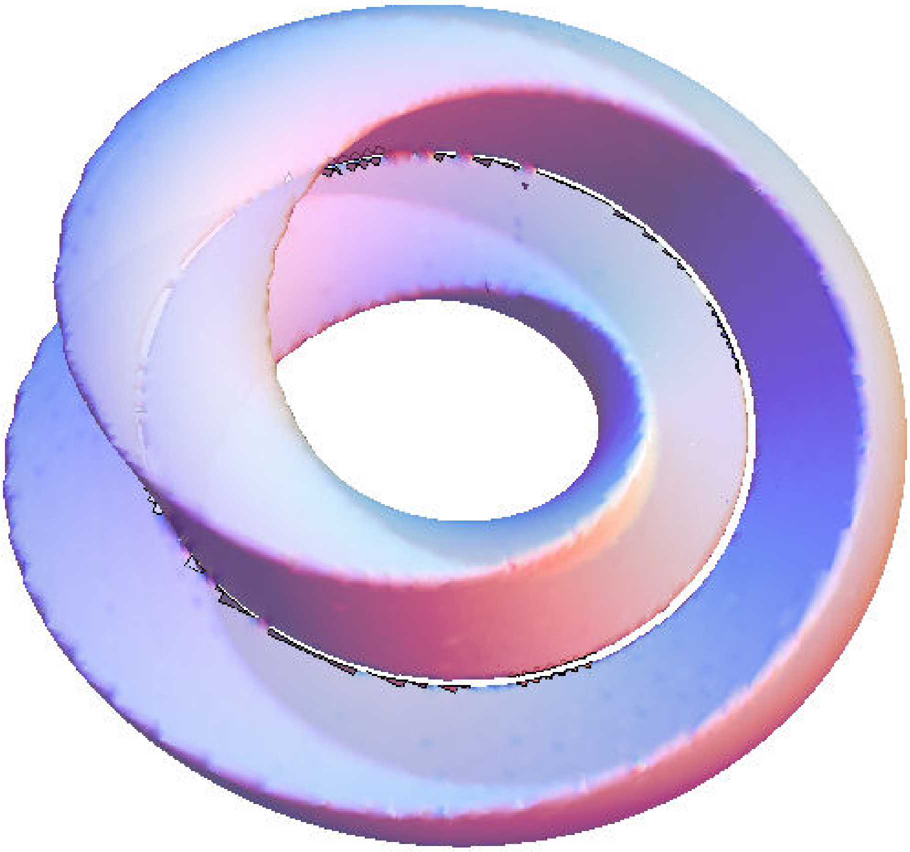


Figure 1. The superlevel set N_p of the toroid discussed in Example 2.2. Courtesy of R. Gagné.

about the z-axis, where $0 < r_1 < r_2$. Its boundary ∂X is a torus. We use *toric coordinates* (r, θ, ϕ) in X, that is, (r, θ) are polar coordinates for D_{r_1,r_2} and ϕ is the angle of revolution. We define $f(r, \theta, \phi)$ by the formula

$$f(r, \theta, \phi) = r^2 \left(1 - 2\sin^2(\theta - \phi/2)\right). \tag{1}$$

Note that at $\phi = 0$, that is, when $y = 0$, we have

$$f(r, \theta, 0) = r^2(1 - 2\sin^2\theta) = (x - r_2)^2 - z^2,$$

so the center of D_{r_1,r_2} is a simple saddle. The revolution of the center of D_{r_1,r_2} produces a circle S of radius r_2 given by $r = 0$. Both f and ∇f vanish on S, so this circle is a singular curve of f forming the self-intersection of the isosurface $f = 0$. Moreover, all points on it are simple saddles of the function f restricted to the plane perpendicular to the circle. When the disc D_{r_1,r_2} is revolved about the angle ϕ, the axes of outward and inward trajectories of its central saddle are revolved about the angle $\phi/2$. Let now N be a neighborhood of S in X defined in toric coordinates by $r \leq \delta$, for any chosen $0 < \delta < r_1$. This is a smaller toroid whose surface is a torus $T := \partial N$ given by the equation $r = \delta$. The set N_p is shown in Figure 1, the set N_n is not shown, because it has the same form. The sets L_p and L_n are the corresponding double-winded bands on T. It is seen that N_p and N_n are connected. If we wanted to use this information only, S would be classified as a regular region. In the next section, we run our classification algorithm on this example and show that our homology criterion classifies it as a saddle-type region.

REMARK 2.3. Although Example 2.2 has been produced with the purpose of illustrating our method, it has been observed that it may provide an interesting insight into the understanding of the atomic orbital of hydrogen in a strong magnetic field. Our function f assumes negative values but we may consider its translation $\rho = f + c$, $c > 0$, interpreted as the probabilistic density of the appearance of the electron in a toroid around the proton.

Below we give a brief overview of the Conley index[?, ?] in the context of a flow generated by the gradient field of a C^2 function f.

A subset $S \subset X$ is called an *invariant set* of a flow φ, if $\varphi(t,x) \in S$ for all $x \in S$ and all $t \in \mathbb{R}$. It is called *isolated* if there exists a neighborhood N such that S is the maximal invariant set in N. The main purpose of the Conley index theory of flows is to describe isolated invariant sets.

Let N be the closure of a bounded region in X and ∂N its boundary. A set $L \in N$ is called the *exit set* of N if given $x \in N$ and $t_0 > 0$ such that $\varphi(t_0, x) \notin N$, then there exists $0 \le t_1 \le t_0$ such that $\varphi(t_1, x) \in L$. Obviously, we must have $L \subset \partial N$. The following well-known result gives rise to the homological Conley index theory:

THEOREM 2.4. *[?, Prop. 10.40] Suppose that L is a closed subset of ∂N. If $H_*(N,L)$ is non-trivial then N contains a non-empty invariant set in its interior.*

The *homological Conley index* of an invariant set S is given by $CH_*(S) = H_*(N,L)$. Additional assumptions on N and L are imposed in order to make the index definition dependent only on S, and not on the choice of N, and in order to achieve desired topological properties, such as stability of the index with respect to small perturbations of the flow. We avoid introducing those assumptions here because, in our work, we will not use the whole extent of the Conley index theory; rather, we use it as an inspiration for the digital setting.

Here are the two main ideas leading to our algorithms sorting critical components.

The first idea is, whether critical points of a function f are isolated or not, they can be grouped into connected components C. Those components may be regarded as invariant sets of the flow associated to the gradient field ∇f. If those components are bounded and isolated in X, the Conley index theory can be applied to analyse their criticality.

The next idea is to extend the described topological analysis directly to functions defined on discrete data, without searching for their smooth interpolation. The anticipation of success comes from the fact that homology is a computable topological invariant much better adapted to combinatorial structures such as cubical grids than to structures coming from smooth analysis.

3. DIGITAL CRITICAL REGIONS

In this section we introduce an analogy of critical points and critical regions for scalar fields defined on discrete subsets of $\mathbb{R}^d$.

We adopt the framework of so-called *cubical grids* $\mathcal{K}$ defined in Kaczynski *et al.*[?] Recall that $\mathcal{K}$ is the collection of *elementary cubes* of the form

$$Q = I_1 \times I_2 \times \cdots \times I_d$$

where $I_j = [k, k+1]$ or $I_j = \{k\}$, $k \in \mathbb{Z}$ (the set of integers), including the empty set. We denote by $\mathcal{K}^d$ the subset of $\mathcal{K}$ consisting of d–dimensional cubes, also called *full* cubes. These are the cubes which have no degenerate intervals $I_j = \{k\}$ in their expression, equivalently, which are not proper faces of other cubes in $\mathcal{K}$. Note that $\mathcal{K}$ is a particular case of a regular cellular complex. A set $X \subset \mathbb{R}^d$ is called a *cubical set* if it is a finite union of elements of $\mathcal{K}$, and it is called *full* if it is a finite union of elements of $\mathcal{K}^d$. We should distinguish between a finite set $\mathcal{X} \subset \mathcal{K}$ which is a combinatorial object and the *carrier* of $\mathcal{X}$ which is the cubical set $X \subset \mathbb{R}^d$ given by

$$X = |\mathcal{X}| = \bigcup \mathcal{X}.$$

Given a cubical set X, $\mathcal{K}(X)$ is the restriction of the grid $\mathcal{K}$ to the cubes contained in X.

Whenever chosen units appear too coarse, one applies the rescaling isomorphism $\Lambda^k : \mathbb{R}^d \to \mathbb{R}^d$,

$$\Lambda^k(x_1, x_2, \ldots, x_d) = (kx_1, kx_2, \ldots, kx_d),$$

where $k \in \mathbb{Z}$ is called a *scaling factor*. The corresponding refined grid is the image of $\mathcal{K}$ under the *inverse rescaling* $\Lambda^{1/k} = (\Lambda^k)^{-1}$. Given $\mathcal{A} \subset \mathcal{K}^d$, the *k-th subdivision* of $\mathcal{A}$ is given by

$$\operatorname{sd}^k \mathcal{A} := \{Q \in \Lambda^{1/k} \mathcal{K}^d \mid Q \subset A\}.$$

Here is the natural extension of topological concepts of neighborhood and boundary to the digital setting.

DEFINITION 3.1. Let A be a bounded set in $\mathbb{R}^d$. The combinatorial *wrap* of A is a subset of $\mathscr{K}$ defined by

$$wrap\,(A) = \{P \in \mathscr{K}^d \mid P \cap A \neq \emptyset\}.$$

It is a finite set, so its carrier denoted by $\mathrm{wrap}\,(A) = |wrap\,(A)|$ is a cubical set. Given a finite set $\mathscr{A} \in \mathscr{K}^d$, its combinatorial *outer boundary* is defined by

$$bd\,(\mathscr{A}) := wrap\,(A) \setminus \mathscr{A}.$$

Its carrier is called the *outer boundary* of $A = |\mathscr{A}|$ and denoted by $\mathrm{bd}\,(A)$. The *factor k scaled wrap* of A is the subset of $\Lambda^{1/k}\mathscr{K}$ defined by

$$wrap^k(A) = \{P \in \Lambda^{1/k}\mathscr{K}^d \mid P \cap A \neq \emptyset\}.$$

Its carrier with respect to the refined grid is denoted by $\mathrm{wrap}^k(A)$. The *factor k scaled outer boundaries* of $\mathscr{A}$, respecively $A = |\mathscr{A}|$, are the sets

$$bd^k(\mathscr{A}) := wrap^k(A) \setminus \mathrm{sd}^k(\mathscr{A}),$$

and, respectively,

$$\mathrm{bd}^k(A) := |bd^k(\mathscr{A})|.$$

The study of scaled wraps is motivated by the following theorem.

THEOREM 3.2. (Allili[8]) *Given a cubical set A and a scaling factor $k \geq 3$, the inclusion $A \hookrightarrow \mathrm{wrap}^k(A)$ induces an isomorphism in homology.*

In practice, the values of functions $f : \mathbb{R}^d \to \mathbb{R}$ are only known at finitely many grid points. Our approach, inspired by combinatorial multivalued maps[?,?] is to define combinatorial maps on full elementary cubes rather that on their vertices. Thus we consider discrete functions

$$f : \mathscr{X}^d \to \mathbb{R}, \text{ where } \mathscr{X}^d \subset \mathscr{K}^d.$$

This approach is very natural if we view the smallest objects in an image, n-pixels, as d-cells of a cubical grid $\mathscr{K}$. The passage from our definition to a vertex definition is by simple translation of coordinates $x_i \mapsto x_i + 1/2$, $i = 1, 2, \ldots d$. Conversely, if a function f is given on a cubical set $X = |\mathscr{X}^d| \subset \mathbb{R}^d$, we may define the discretization of f on $\mathscr{X}^d$ by taking its values at the center of each elementary full cube.

DEFINITION 3.3. Let $\mathscr{X}^d \subset \mathscr{K}^d$. Consider a function $f : \mathscr{X}^d \to \mathbb{R}$ and an elementary cube $Q \in \mathscr{X}^d$ such that $wrap\,(Q) \subset \mathscr{X}^d$. The *upper wrap* of Q is defined by

$$\overline{wrap}\,(Q) := \{P \in \; wrap\,(Q) \mid f(P) > f(Q)\}\,.$$

Analogously, the *lower wrap* of Q is given by

$$\underline{wrap}\,(Q) := \{P \in \; wrap\,(Q) \mid f(P) < f(Q)\}\,,$$

and the *level wrap* of Q is given by

$$wrap_z(Q) := \{P \in \; wrap\,(Q) \mid f(P) = f(Q)\}\,.$$

The extension of the upper and lower wraps to cubical sets $A = |\mathscr{A}|$ is more delicate. We assume that $wrap\,(A) \subset \mathscr{X}$. Since the values of f may vary, we want the inequalities in the earlier formulas to be satisfied locally:

$$\overline{wrap}\,(A) := \{P \in bd\,(\mathscr{A}) \mid f(P) > f(Q) \text{ for all } Q \in \mathscr{A} \cap bd\,(P)\}\,,$$

$$\underline{wrap}\,(A) := \{P \in bd\,(\mathscr{A}) \mid f(P) < f(Q) \text{ for all } Q \in A \cap bd\,(P)\}\,.$$

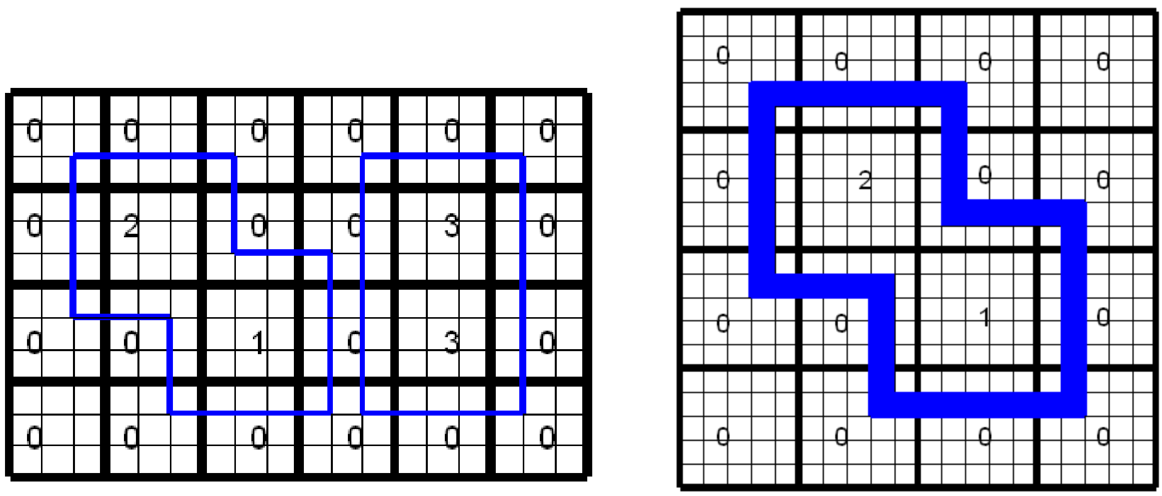
$$wrap_z(A) := bd\,(\mathscr{A}) \setminus \left(\underline{wrap}\,(A) \cup \overline{wrap}\,(A)\right)$$

As for wraps, the notation $\overline{\mathrm{wrap}}$ and $\underline{\mathrm{wrap}}$ is used for the carriers of upper and lower wraps $\overline{wrap}$ and $\underline{wrap}$.

Analogous terminology and notation is carried over to the scaled wraps introduced in definition 3.1.

The upper and lower wraps are analogues to the exit set and entrance set from the Conley index theory. These sets are used by Allili *et al.*,[?] where the algorithms detecting critical pixels in $\mathbb{R}^2$ and their critical components $\mathscr{C} \subset \mathscr{K}^2$ are based on counting the number of 4–connected and 8–connected components of upper and lower wraps $\overline{wrap}(\mathscr{C})$ and $\underline{wrap}(\mathscr{C})$. In this setting, scaled wraps introduced above are not necessary but, as we showed in the previous section, the connectivity concept is not sufficient in dimensions higher than 2. In order to compute homology groups, we need to pass to geometric carriers. Therefore, if we follow the approach chosen in Allili *et al.*,[?] the set $N = \mathrm{wrap}(C)$ would be a natural candidate for the isolating neighborhood of $\mathscr{C} \subset \mathscr{K}^2$, and $\overline{\mathrm{wrap}}(\mathscr{C})$ and $\underline{\mathrm{wrap}}(\mathscr{C})$ respectively, for its entrance and exit sets. However, there is a problem related to the fact that, while the combinatorial sets $\overline{wrap}(\mathscr{C})$ and $\underline{wrap}(\mathscr{C})$ are disjoint, their carriers may intersect, resulting in misleading topological information. One possible remedy to the described problem is to consider open wraps instead of the closed ones. However, computing homology of open polyhedra is a substantially more complex task than that of compact polyhedra. We shall achieve this effect, while staying in the class of compact cubical sets, by considering "two-layer factor 5 scaled wraps" as seen in Figure 2. This is explained in the following example.

EXAMPLE 3.4. The two images in Figure 2 show a maximal component $\mathscr{C}$ of two pixels with the scalar values 1 and 2. The displayed set L_n follows the whole boundary of N, the candidate for the isolating neighborhood of $\mathscr{C}$. If we used the unit-scale boundaries in the definition of N, the pixels of L_n would intersect at the common vertex of the two pixels of $\mathscr{C}$, so L_n would contain two cycles and provide false topology information. Using the scaling factor $k = 3$, as seen in (a), requires the use of a thin boundary in order to ensure that the boundaries of the isolating neighborhoods of separate critical components do not intersect each other. For technical reasons, we prefer using only full cubes, and therefore require a scaling factor $k = 5$ in order to use a thick boundary, as seen in (b).



(a) (b)

Figure 2. Two choices of the sets N and L_n for the maximal component consisting of pixels with scalar values 1 and 2, discussed in Example 3.4. The scaling factor $k = 3$ is used in (a) and $k = 5$ in (b).

DEFINITION 3.5. Let $\mathscr{X}^d \subset \mathscr{K}^d$ and $k \geq 5$. Consider a function $f : \mathscr{X} \to \mathbb{R}$ and an elementary cube $Q \in \mathscr{X}^d$ such that $wrap(Q) \subset \mathscr{X}^d$. Define

$$\begin{aligned} N &:= \mathrm{wrap}^k(\mathrm{wrap}^k(Q)), \\ L_p &:= \overline{\mathrm{wrap}}(Q) \cap \mathrm{bd}^k(\mathrm{wrap}^k(Q)), \\ L_n &:= \underline{\mathrm{wrap}}(Q) \cap \mathrm{bd}^k(\mathrm{wrap}^k(Q)), \\ L_z &:= \mathrm{wrap}_z(Q) \cap \mathrm{bd}^k(\mathrm{wrap}^k(Q)). \end{aligned}$$

The elementary cube Q is called *ordinary* if $H_*(N, L_p) = 0$ and $H_*(N, L_n) = 0$. Otherwise, it is called *singular*.

Here is the related algorithm.

ALGORITHM 3.6. Detecting singular cubes

For each elementary full cube Q
build N, L_p, L_n
$H := H_*(N, L_p) = 0$ *and* $H_*(N, L_n) = 0$
    if $H =$ *TRUE* **then** Q *is ordinary*
        else Q *is singular*
    endif

As it is pointed out by Allili *et al.*,? the adjacent singular cells may, in a sense, cancel each other. In other words, the type of criticality of a singular cube cannot be decided without looking at neighboring cubes. This is taken care of by the following definitions.

DEFINITION 3.7. Let $\mathscr{X}^d \subset \mathscr{K}^d$, $k \geq 5$ and consider a function $f : \mathscr{X}^d \to \mathbb{R}$. A set $\mathscr{C}$ of singular elementary cubes in $\mathscr{K}^d$ is called an *isolated singular component* if

(a) $\operatorname{wrap}(C) \subset X$;

(b) $C = |\mathscr{C}|$ is connected;

(c) $\mathscr{C}$ is *isolated* in the sense that any $P \in bd(\mathscr{C})$ is ordinary.

The *isolating neighborhood* of $C = |\mathscr{C}|$, its *lower*, *upper*, and *level sets* are respectively defined by

$$\begin{aligned} N &:= \operatorname{wrap}^k(\operatorname{wrap}^k(C)), \\ L_p &:= \overline{\operatorname{wrap}}(Q) \cap \operatorname{bd}^k(\operatorname{wrap}^k(C)), \\ L_n &:= \underline{\operatorname{wrap}}(Q) \cap \operatorname{bd}^k(\operatorname{wrap}^k(C)), \\ L_z &:= \operatorname{wrap}_z(Q) \cap \operatorname{bd}^k(\operatorname{wrap}^k(C)). \end{aligned}$$

DEFINITION 3.8. An isolated singular component $\mathscr{C}$ is called *regular* if $H_*(N, L_p) = 0$ and $H_*(N, L_n) = 0$. Otherwise it is called a *critical component*. A singular cube Q is called *critical*, if it belongs to a critical component, otherwise, it is called *regular*. An ordinary cube Q is, by definition, *regular*.

In a sense, singular-regular components are an analogy of inessential or removable singularities in mathematical analysis.

Given any singular cell $Q \subset \mathscr{K}^d$, the component $\mathscr{C}(Q) = \mathscr{C}$ is constructed by the following algorithm:

ALGORITHM 3.9. Sorting components in $\mathbb{R}^d$

For each singular cube Q, $\mathscr{C} := \{Q\}$
while $P \in bd(\mathscr{C}) \cap X$ *is singular*
    $\mathscr{C} := \mathscr{C} \cup \{P\}$
    build $bd(\mathscr{C})$
endwhile
build N, L_p, L_n, L_z
do
    $H := H_*(N, L_p) = 0$ *and* $H_*(N, L_n) = 0$
        if $H =$ *TRUE* **then** $\mathscr{C}$ *is a regular component*
        else if $L_p = L_z = \emptyset$ **then** $\mathscr{C}$ *is a maximum component*
        else if $L_n = L_z = \emptyset$ **then** $\mathscr{C}$ *is a minimum component*
        else $\mathscr{C}$ *is a saddle component*
    endif
break

Our algorithms are implemented in a C++ program which makes use of the software CHomP? for relative homology computations. We discuss below examples of experimentation in $\mathbb{R}^2$ and $\mathbb{R}^3$.

EXAMPLE 3.10. We have run the program for the function $f(x,y) = (x^2+y^2-5)^2$ on the domain $X = [-3,3]^2$, discretized on a planar grid of 51^2 2D–cubes, with the function's range quantized on 2048 levels. The results are displayed in Figure 3. It is worth noting that discretizing the range is necessary when the domain is discrete. This is done so far in a heuristic way. The four saddle pixels and four maxima on the boundary of the domain appear due to forcing the global minimum of f outside of X. The aim of this forcing is to simulate the topology of a sphere, allowing the eventual use of the Euler number as a correctness check.

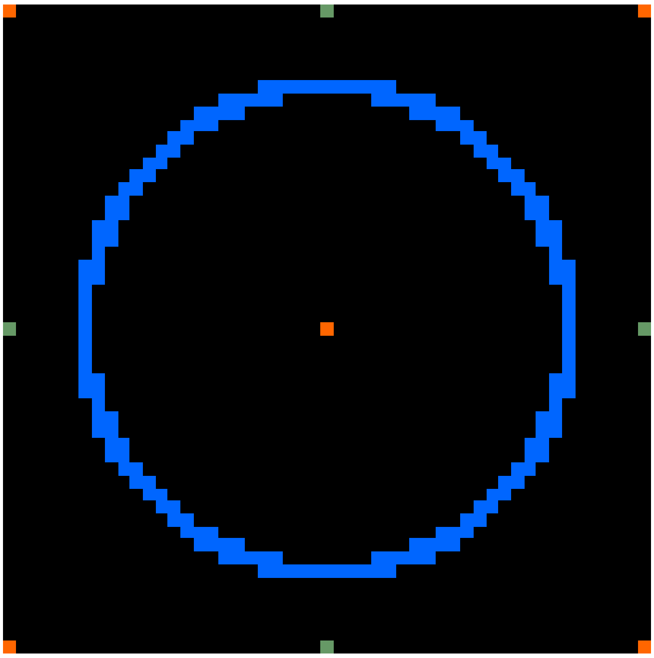


Figure 3. Critical components of the function $f(x,y) = (x^2+y^2-5)^2$, discretized as discussed in Example 3.10: the minimal circle, a maximum pixel in the center, four saddle pixels and four maxima on the boundary of the domain.

EXAMPLE 3.11. We have also run the program for the function $f(x,y,z) = (x^2+y^2+z^2-5)^2$, on the domain $X = [-3,3]^3$, discretized on a spatial grid of 51^3 cubes, with the range quantized on 8192 levels. Figure 4 shows the isolating neighborhood N of the minimal sphere and the set L_p which consists of two layers of surrounding spheres. This set is hollow on the inside.

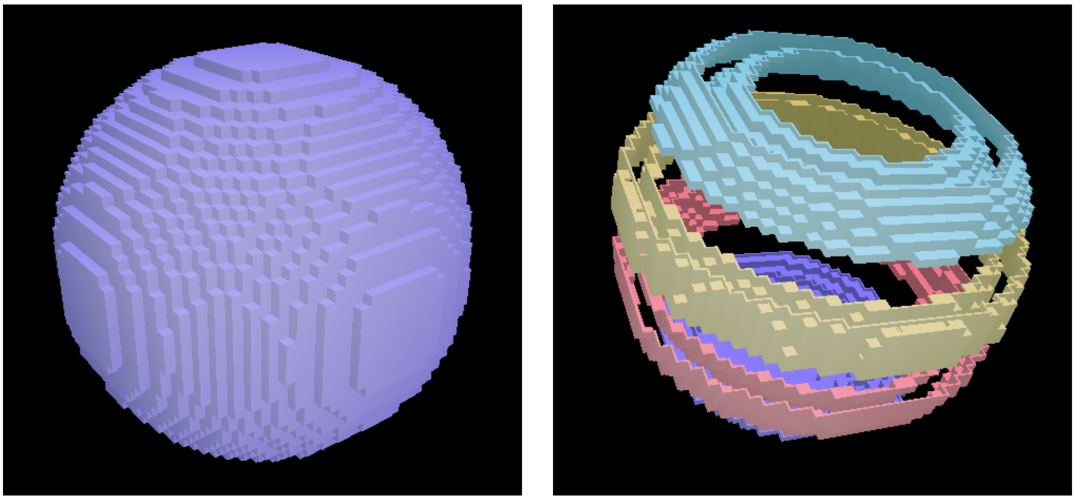
EXAMPLE 3.12. Our program applied to Example 2.2 produces a toroid N and the sets L_p and L_n which are surfaces of the set displayed in Figure 1's exterior with respect to N. In particular, we have obtained the following homology results:

$$H_*(N,L_p) \simeq H_*(N,L_n) \simeq \begin{cases} \mathbb{Z} & \text{if } k=2 \\ \mathbb{Z} & \text{if } k=1 \\ 0 & \text{if } k=0 \end{cases}.$$

N is shown in Figure 3 (a) and L_p, in Figure 3 (b). L_n is not shown, being similar to L_p. We have used the domain $X = [-3,3]^2 \times [-1/5,1/5]$ discretized on a spatial grid of $101^2 \times 7$ cubes, with the range quantized over 16384 levels.

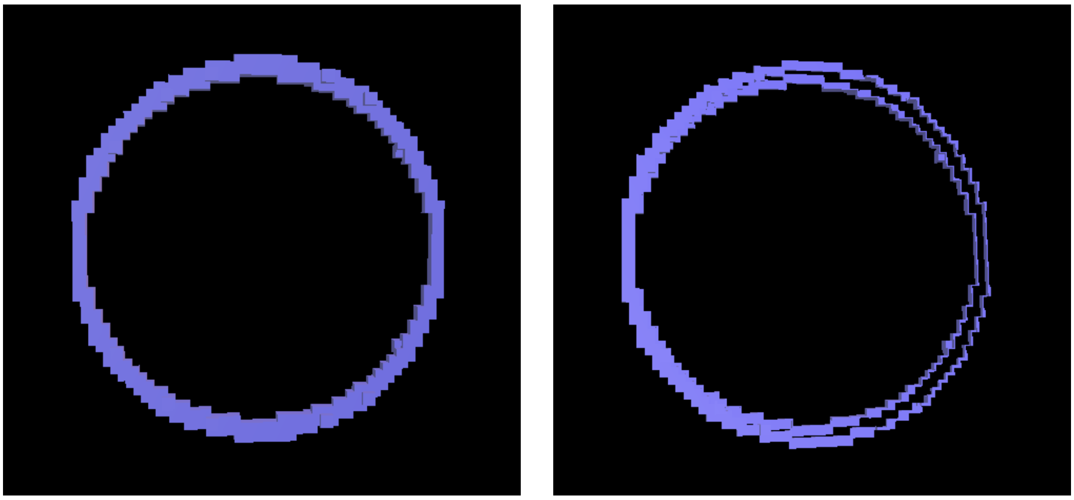
We finish this section with several remarks.

REMARK 3.13. In the definition of the homological Conley index for flows, the condition $H_*(N,L_p) \neq 0$ would be sufficient for criticality, and one would not need to take N_n into account. This is due to the fact, that the flow of $-\nabla f$ has



(a) (b)

Figure 4. The set N in (a) and several slices of the set L_p in (b) for the minimal sphere discussed in Example 3.11. The whole set L_p consists of two concentric spheres. The colors (online version) are used to facilitate viewing.



(a) (b)

Figure 5. The sets N in (a) and L_p in (b) for the function first described in Example 2.2, as described in Example 3.12.

the same trajectories as that of ∇f oriented in the opposite direction. In the digital case we study here, there is no such symmetry, so we need to take both L_p and L_n into consideration.

REMARK 3.14. The condition $H^*(N,L_p) = 0$ (and the same for L_n) in our definitions and algorithms can be replaced by the stronger condition that the homotopy type $[N,L_p]$ of the pair of spaces (N,L_p) be non-trivial. However, as already commented in the introduction, such a condition could not be verified in practice. A yet stronger but combinatorial condition would be "N collapses to L_p", which means that N can be deformed to N_p by a chain of *elementary collapses* described in Ref. **?**, Sec. 2.4. Elementary collapses preserve not only homology but also homotopy type and the algorithm *Collapse* actually is the fastest method of verifying that $H_*(X,L_p) = 0$.

REMARK 3.15. Note that there are many topologically distinct types of saddle components. In the case of a single elementary cube, one can characterize it as a minimum, maximum, or a k–fold saddle, $k \in \mathbb{N}$, by means of $H_*(N,L_p)$ and $H_*(N,L_p)$. For larger components $\mathscr{C}$, the topology of C may be non-trivial, as we have seen it in examples presented earlier in this section. Thus one has to also take into account $H_*(C)$.

4. DIRECTIONS OF THE FUTURE WORK

A natural question to raise is whether or not our construction is stable with respect to perturbation of f or with respect to rescaling (that is, with respect to a change of resolution). Since our model assumes a discrete function as an input, and stability concerns continuous data, these questions cannot be answered within the framework of this paper, because that would involve different methods of discretizing, that is interpolating a continuous fonction. Thus interpolation methods and multiresolution techniques[?] should be integrated into the study. Also, the concept of *topological persistence* due to Edelsbrunner, Zomorodian *et al.*[?,?] may become a helpful tool in answering the stability questions.

Another important direction is an application of our results to the problem of isosurfaces which has been the main motivation for our work as commented in the introduction. Note that we apply the term *isosurface* inclusively, for level sets of scalar fields of an arbitrary number of variables. In particular, we are interested in studying isovolumes.

Yet another interesting direction is to study extensions of the Euler formula in the context of our work. This formula is used (sometimes enforced) as a criterion of correctness of critical points extracted from discrete data. Our two-dimensional experiments test positively when the isolation and non-degeneracy assumptions hold and when there are no critical pixels on the boundary of the region. But the Euler formula cannot hold in the presence of critical components. In fact, a critical component itself may have an Euler characteristic different from that of a point. Thus the goal is to find a suitable generalization of the Euler formula that would hold for the degenerate case.

ACKNOWLEDGMENTS

M. Allili and T. Kaczynski are supported by the NSERC of Canada discovery grant.

REFERENCES

1. G. H. Weber, G. Scheuermann, and B. Hamann, "Detecting critical regions in scalar fields," *IEEE TCVG Symposium on Visualization* , pp. 1–11, 2003.
2. Y. Matsumoto, *An Introduction to Morse Theory*, vol. 208 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, 2002.
3. J. D. Boissonnat and M. T. (editors), *Effective Computational Geometry for Curves and Surfaces*, Math. and Visualization Series, , Springer-Verlag, 2007.
4. H. Edelsbrunner, J. Harer, and A. Zomorodian, "Hierarchical morse-smale complexes for piecewise linear 2-manifolds," *AMC Symposium on Computational Geometry, Discrete & Compu. Geom.* **30 1**, pp. 81–107, 2003.
5. H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, "Morse–smale complexes for piecewise linear 3-manifolds," *Proc. IEEE Conf. Visualization* , pp. 275–280, 2004.
6. V. I. Arnold, *Catastrophe Theory*, Springer-Verlag, 1984.
7. J. Milnor, *Singular Points of Complex Hypersurfaces*, vol. 61, Princeton Univ. Press, Annals of Mathematics Studies, 1968.
8. M. Allili, "Une approche algorithmique pour le calcul de l'homologie de fonctions continues," *Ph. D. Thesis, Dep. de math. et d'info., Univ. de Sherbrooke* , 1998.